

Szegedi Tudományegyetem
Informatikai Tanszékcsoporth

Affin deformációk paramétereinek becslése megfeletetések nélkül

PhD-értekezés tézisei

Domokos Csaba

Témavezető:

Dr. Kató Zoltán

Szeged
2010

Bevezetés

A *képregisztráció* majdnem minden képfeldolgozási alkalmazás fontos lépése, ahol az objektumok eltérő nézőpontból vagy különböző szenzorral készített képeit kell összehasonlítani esetleg kombinálni. Tipikus felhasználási területei közé tartozik az alakfelismerés, objektumok követése videószekvenciákon, műholdképek változásainak detektálása, szuperfelbontású képek, képmozaikozás, illetve az orvosi képfeldolgozás. Általánosan a feladat egy olyan transzformáció megkeresése, mely az egyik képet (*megfigyelés*) a második képhez illeszti (*sablon*). A lehetséges transzformációk nagy száma miatt a probléma *rosszul definiált*, hacsak nem vesszük figyelembe ezt a nagy változatosságot.

Sok esetben a képjellemzők változása annyira összetett, hogy az egyetlen járható út a képek regisztrálására, hogy a bináris reprezentációjuk alapján oldjuk meg a problémát. A röntgen képek (lásd 1. ábra) esete jó példa arra, hogy egy erősen nemlineáris színtorzulás megnehezíti a regisztrációt. Ha a képek tökéletes szürkeárnyalatos változata adott lenne, akkor az illesztő transzformáció paramétereinek becslése visszavezethető lenne egy lineáris egyenletrendszer megoldására [1; 2]. Mindazonáltal valós alkalmazások során ilyen erős feltétel ritkán teljesül. Itt egy olyan regisztrációs eljárást mutatunk be, amely a kép intenzitásértékeinek felhasználása nélkül oldja meg a feladatot. A képek között valós esetben projektív transzformáció van (melyet síkhomográfiának is hívnak). Bár a projektív transzformációk nemlineárisak, gyakran affin transzformációkkal jól modellezhetők, melyek viszont már lineárisak. Emiatt az affin transzformációk fontos szerepet játszanak a képregisztrációban.

Ez a tézis a Szerző azon eredményeit foglalja össze, melyek a bináris képregisztráció pontmegfeleltetések nélküli megoldásával kapcsolatosak, és ahol affin transzformációkat feltételezünk a regisztráció során.

Alapmegoldás

Először bevezetünk néhány szükséges jelölést, majd formalizáljuk a megoldás alapötletét. Jelölje az n -dimenziós *sablon* és *megfigyelés* képek homogén pontkoordinátáit $\mathbf{x} = [x_1, x_2, \dots, x_n, 1]^T$ és $\mathbf{y} = [y_1, y_2, \dots, y_n, 1]^T \in \mathbb{P}^n$. Az alakzatok között fennálló reláció a következő [3; 4]

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y}, \quad (1)$$

ahol \mathbf{A} egy ismeretlen affin transzformáció, melyet meg szeretnénk határozni:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} & a_{1(n+1)} \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & a_{n(n+1)} \\ 0 & \dots & 0 & 1 \end{bmatrix}, \quad \text{és} \quad \mathbf{A}^{-1} = \begin{bmatrix} q_{11} & \dots & q_{1n} & q_{1(n+1)} \\ \vdots & \ddots & \vdots & \vdots \\ q_{n1} & \dots & q_{nn} & q_{n(n+1)} \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

Egy olyan *direkt megoldást* szeretnénk találni, mely során nem kell pontmegfeleltetéseket keresni.

Ha olyan képi jellemzők is rendelkezésre állnak, melyek invariánsak a transzformációra nézve (például a pixelek intenzitásértékei [2]), akkor további relációkat definiálhatunk:

$$f(\mathbf{x}) = g(\mathbf{Ax}) = g(\mathbf{y}), \quad (2)$$

ahol $f, g : \mathbb{P}^n \rightarrow \mathbb{R}$ *kovariánsak* az \mathbf{A} transzformációra nézve. A fenti összefüggések továbbra is érvényben maradnak, ha egy függvény hat az (1) [2; 3; 4] vagy a (2) egyenlet [5; 6; 7] mindkét oldalán. Valóban, tetszőlegesen megválasztott $\omega_p : \mathbb{P}^n \rightarrow \mathbb{R}$ és $\omega_c : \mathbb{R} \rightarrow \mathbb{R}$ függvények esetén kapjuk, hogy

$$\omega_p(\mathbf{y}) = \omega_p(\mathbf{Ax}), \quad \text{és} \quad (3)$$

$$\omega_c(g(\mathbf{y})) = \omega_c(g(\mathbf{Ax})) = \omega_c(f(\mathbf{x})). \quad (4)$$

Így ezen nemlineáris ω_p és ω_c függvények felhasználásával annyi lineárisan független egyenletet generálhatunk, amennyire szükségünk van. A nemlineáris ω_p függvény közvetlenül a pontkoordinátákon hat, így az ismeretlen \mathbf{A} transzformáció paraméterein, és egy *nemlineáris* egyenletrendszerhez vezet [3; 4]. Ezzel szemben ω_c az f és g *kovariáns függvényeken* hat, és alkalmazásával egy *lineáris* egyenletrendszert kapunk [2; 5; 6]. Az alakzatokat az $\mathbb{1} : \mathbb{P}^n \rightarrow \{0, 1\}$ karakterisztikus függvénnyel reprezentáljuk, ahol 0 és 1 a háttérhez, illetve az előtérhez tartozó pixeleket jelöli. Ha a *sablon képet* $\mathbb{1}_t$ -vel, a *megfigyelést* pedig $\mathbb{1}_o$ -val jelöljük, akkor az (1) egyenletből a következő összefüggéshez jutunk [3; 4]

$$\mathbb{1}_t(\mathbf{x}) = \mathbb{1}_o(\mathbf{Ax}) = \mathbb{1}_o(\mathbf{y}). \quad (5)$$

Bináris képek affin torzulásinak paramétereinek becslése: Polinomiális megoldás

Most egy olyan újfajta megközelítést mutatunk be, ahol a pontos transzformáció, mely egy ismert 2D alakzat és annak egy deformált *megfigyelése* között van, egy polinom egyenletrendszer megoldásával megkapható. A klasszikus *képjellemzőkön alapuló* megközelítések először pontpárokat azonosítanak a képeken, majd az (1) egyenletben megadott lineáris egyenletrendszert oldják meg. Ellenben mi egy *direkt megoldást* szeretnénk találni a pontmegfeleltetés problémájának megoldása nélkül.

Az (1) egyenletben megadott relációból kiindulva, majd véve mindkét oldal Lebesgue integrálját kapjuk, hogy [3; 4]

$$\int_{\mathbb{P}^n} \mathbf{x} \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathbb{P}^n} \mathbf{A}^{-1} \mathbf{y} \, d\mathbf{y} , \quad (6)$$

ahol az $\mathbf{x} = \mathbf{A}^{-1} \mathbf{y}$, $d\mathbf{x} = |\mathbf{A}^{-1}| d\mathbf{y}$ integrál transzformációt alkalmaztuk. Az $|\mathbf{A}|$ az ún. Jacobi-féle determináns, melyről feltesszük, hogy mindig pozitív. Továbbá kiszámolható az alábbi integrálokkal [3; 4]

$$\int_{\mathbb{P}^n} \mathbb{1}_t(\mathbf{x}) \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathbb{P}^n} \mathbb{1}_o(\mathbf{y}) \, d\mathbf{y} \quad \Leftrightarrow \quad |\mathbf{A}| = \frac{\int_{\mathbb{P}^n} \mathbb{1}_o(\mathbf{y}) \, d\mathbf{y}}{\int_{\mathbb{P}^n} \mathbb{1}_t(\mathbf{x}) \, d\mathbf{x}} ,$$

melyek közvetlenül számolhatók az input képek alapján. Mivel a karketrisztikus függvények csak 0 és 1 értéket vesznek fel, a fenti integrált tovább egyszerűsíthetjük [3; 4]:

$$\int_{\mathbb{P}^n} \mathbb{1}_t(\mathbf{x}) \, d\mathbf{x} \equiv \int_{\mathcal{F}_t} d\mathbf{x} ,$$

ahol $\mathcal{F}_t = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{1}_t(\mathbf{x}) = 1\}$ az integrálási tartományt, a *sablon* kép (véges) előtér régióját jelöli ki. A *megfigyelés* esetén hasonlóan, $\mathbb{1}_o(\mathbf{y})$ integrálásakor is megszoríthatjuk a tartományt az \mathcal{F}_o előtér régióra. Összeszorozva egymással a (6) és az (5) egyenletet, az alábbi, véges tartományon vett integrál egyenletet kapjuk [3; 4]:

$$\int_{\mathcal{F}_t} \mathbf{x} \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathcal{F}_o} \mathbf{A}^{-1} \mathbf{y} \, d\mathbf{y} . \quad (7)$$

Ebből az következik, hogy \mathcal{F}_t és \mathcal{F}_o között is fenáll az $\mathcal{F}_o = \mathbf{A} \mathcal{F}_t$ reláció. Tulajdonképpen az egész alakzatot illesztjük, ahelyett hogy az egyes pontok közötti megfeleltetéseket vennénk. A (3) egyenlet felhasználásával (7)-ből az alábbi integrál egyenlethez jutunk [3; 4]

$$\int_{\mathcal{F}_t} \omega(\mathbf{x}) \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathcal{F}_o} \omega(\mathbf{A}^{-1} \mathbf{y}) \, d\mathbf{y} . \quad (8)$$

Az alapötlet ezen megközelítés mögött az, hogy lineárisan független ω függvények felhasználásával elegendő számú lineáris egyenletet generálunk. Megjegyezzük, hogy az így kapott egyenletek nem tartalmaznak új információt, csupán újabb lineárisan független megszorításokat eredményeznek.

Állítás 1.1 Legyen $\omega : \mathbb{P}^n \rightarrow \mathbb{P}^n$ és $\mathbf{x} \in \mathbb{P}^n$, ahol $n \in \mathbb{N}$. Ha $\omega^{(k)}(\mathbf{x}) = p_k$ jelöli az $\omega(\mathbf{x})$ k -edik koordinátáját, mely egy n -változós valós együtthatós polinom $1 \leq k \leq n$ -re, akkor ω -t a (8) egyenletre alkalmazva egy legfeljebb $\deg(p_k)$ fokú polinom egyenletrendszerrel kapunk.

Mostantól a 2-dimenziós esettel foglalkozunk. Az ω -ra nézve az x^l ($l \in \mathbb{N}_0$) függvényosztály ideális választás. Ekkor a következő polinom egyenleteket kapjuk $k = 1, 2$ -re [3; 4]:

$$|\mathbf{A}| \int x_k^l d\mathbf{x} = \sum_{i=1}^l \binom{l}{i} \sum_{j=0}^i \binom{i}{j} q_{k1}^{l-i} q_{k2}^{i-j} q_{k3}^j \int y_1^{l-i} y_2^{i-j} d\mathbf{y} \quad l = 1, 2, 3. \quad (9)$$

Habár az egyenleteket folytonos tartományon definiáltuk, az integrálokat a gyakorlatban csak közelítően, diszkrét összegek segítségével kapjuk. A képeken elegendő egyszer végigmenni és ezalatt az integrálok, illetve a Jacobi-féle determináns is kiszámítható. Jól látszik, hogy a megoldás egyetlen lépésben, iterációs lépés vagy optimalizálás nélkül megkapható.

Tekintsük most azt az esetet, mikor a megfigyelt pontok koordinátái nem pontosan az eredeti helyükön, hanem csak azok környékén vannak. Ekkor normális eloszlású, additív zajmodellt tekinthetünk a *megfigyelés* pontjainak koordinátáin. Ekkor az (1) egyenlet az alábbi formában írható [4]

$$\mathbf{y}^* = \mathbf{y} + \varepsilon(\mathbf{y}) = \mathbf{A}\mathbf{x} + \varepsilon^*(\mathbf{y}^*) \Leftrightarrow \mathbf{x} = \mathbf{A}^{-1}(\mathbf{y}^* - \varepsilon^*(\mathbf{y}^*)) ,$$

ahol $\varepsilon(\mathbf{y}) \equiv \varepsilon^*(\mathbf{y}^*) = [\varepsilon_1^*(\mathbf{y}^*), \varepsilon_2^*(\mathbf{y}^*), 0]^T$ a zajfüggvény, mely minden $\mathbf{y}^* = [y_1^*, y_2^*, 1]^T$ pontban egy véletlen eltolást eredményez. Összefoglalva az alábbi táblázat mutatja, hogy mekkora a hiba, amit a σ_1, σ_2 szórási paraméterekkel rendelkező normális eloszlású, additív zaj a *megfigyelés* pontkoordinátáin eredményez [4]:

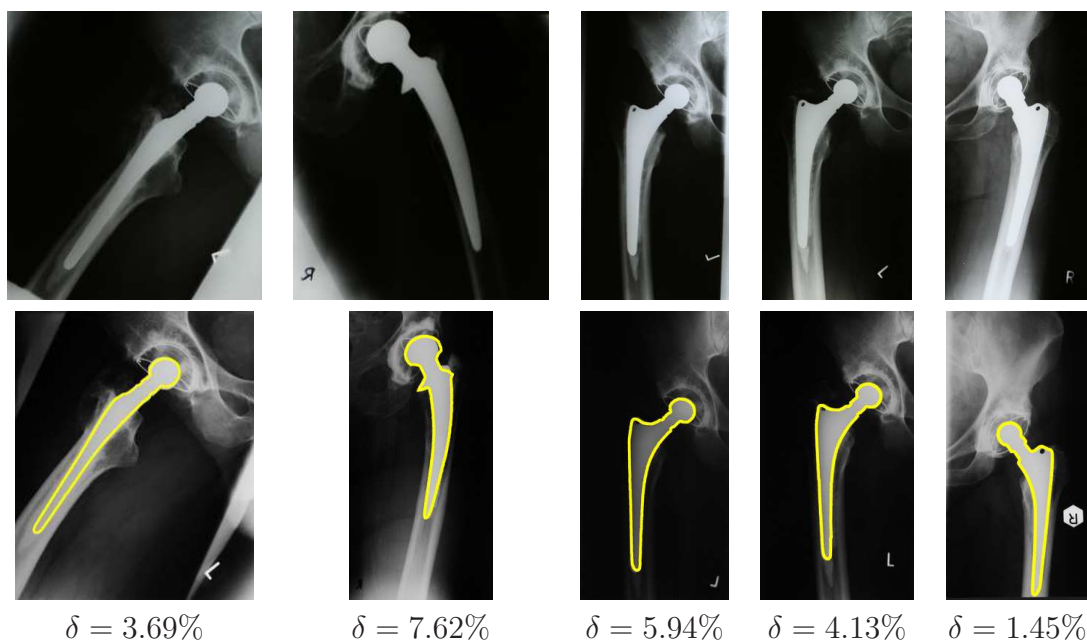
Egyenlet	Hibatag
$ \mathbf{A} $:	0
$\omega(\mathbf{x}) = \mathbf{x}$:	0
$\omega(\mathbf{x}) = [x_1^2, x_2^2, 1]^T$:	$q_{k1}^2 \sigma_1^2 + q_{k2}^2 \sigma_2^2$
$\omega(\mathbf{x}) = [x_1^3, x_2^3, 1]^T$:	$3q_{k3}(q_{k1}^2 \sigma_1^2 + q_{k2}^2 \sigma_2^2)$

A javasolt algoritmust nagy adatbázison teszteltük, mely átlagosan 1000×1000 -es méretű bináris képeket tartalmaz. Az adatbázis összesen 50 000 képből áll, ahol a képek 56 különböző

alakzat és azok transzformált változataiból állnak elő. A regisztráció eredményének kiértékelése céljából, a következő két hibamértéket definiáltuk:

$$\epsilon = \frac{1}{|F_t|} \sum_{\mathbf{p} \in F_t} \|(\mathbf{A} - \hat{\mathbf{A}})\mathbf{p}\|, \quad \text{and} \quad \delta = \frac{|F_r \triangle F_o|}{|F_r| + |F_o|} \cdot 100\%, \quad (10)$$

ahol \triangle a szimmetrikus különbséget, F_t , F_o és F_r pedig a *sablon*, a *megfigyelés* és a *regisztrált* képek előtérpixeleinek halmazát jelöli. Azt is megvizsgáltuk, hogy a javasolt módszer hiányos objektumok esetén mennyire robusztus. Az eredmények mutatják, hogy a módszerünk meglehetősen robusztus, illetve a regisztráció hibája kisebb mértékben nő, mint hasonló módszerek esetében.



1. ábra. Csípőprotézises röntgen képek regisztrációja. Az első sorban lévő képek kontúrjának illesztett változata van rávetítve a második sorban lévő képekre.

Alapvető különbség a klasszikus regisztrációs módszerekhez képest, hogy a mi modellünk jellemző pontok vagy képjellemzők detektálása, illetve optimalizálás nélkül működik, és egy új ötletet felhasználva, a transzformációt egy polinom egyenletrendszer megoldásával határozza meg. A módszer az input képeken elérhető összes információt felhasználja és a transzformáció nagyságától függetlenül megadja a pontos megoldást. A futási eredmények mutatják, hogy a javasolt módszer jó illesztést szolgáltat mind valós, mind pedig szintetikus képek esetén (lásd 1. táblázat). Ezenfelül a módszer zajos *megfigyelések* esetén is robusztus. A módszert sikeresen alkalmaztuk csípőprotézises röntgen képek regisztrációjára (lásd 1. ábra).

1. táblázat. Regisztrációs eredmények 56 alakzat 49 282 megfigyelése esetén. Az esetek 5.47%-ban nem kaptunk megoldást.

	ϵ (pixel)	δ (%)	CPU idő (mp.)
Medián	0.51	0.15	0.98
Középérték	36.98	3.36	0.94
Szórásnégyzet	154.18	12.55	0.2

Alakzatok affin illesztése kovariáns Gauss sűrűségfüggvényekkel

Most egy olyan újfajta megközelítést mutatunk be 2D affin transzformációk becslésére, ahol a transzformációt egy lineáris egyenletrendszer legkisebb négyzetes értelemben vett megoldásával kapjuk meg. Ezt úgy érzük el, hogy Gauss sűrűségfüggvényeket illesztünk az alakzatokra, melyek megőrzik az ismeretlen transzformáció hatását. A javasolt módszer kritikus lépése olyan *kovariáns függvénpárok* megkonstruálása, melyek kielégítik a (2) egyenletet. Bináris képeken ilyen függvények meghatározása nagy kihívás, mivel a képek nem tartalmaznak színinformációt.

Most inhomogén koordinátákat használunk, azaz $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ és $\mathbf{y} = [y_1, y_2, \dots, y_n]^T \in \mathbb{R}^n$. Így az (1)-ben megadott összefüggés az alábbi formában írható [5; 6; 7]

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{t} \Leftrightarrow \mathbf{x} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t}) = \mathbf{A}^{-1}\mathbf{y} - \mathbf{A}^{-1}\mathbf{t}, \quad (11)$$

ahol $(\mathbf{A}, \mathbf{t}) \in (\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1})$ az ismeretlen affin transzformáció, amit meg szeretnénk határozni. A *sablon* kép pontjait tekinthetjük úgy, mint egy $X \sim N(\mu, \Sigma)$ kétdimenziós normális eloszlású véletlen valószínűségű változó mintavételezését, a következő sűrűségfüggvénnyel [5; 6; 7]:

$$p(\mathbf{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right).$$

A továbbiakban a normális eloszlás sűrűségfüggvényeit használjuk, melyek *kovariáns függvénpárokat* határoznak meg az alakzatok felett. Megjegyezzük, hogy a Gauss sűrűségfüggvényekkel nem reprezentáljuk vagy modellezzük az alakzatot, itt csupán a pontkoordináták középértékét és kovarianciáját számoljuk ki. Tetszőleges lineáris transzformációt alkalmazva X -re egy $Y = \mathbf{A}X + \mathbf{t}$ olyan véletlen valószínűségi változót kapunk, mely szintén normális eloszlású, és $X \xrightarrow{(\mathbf{A}, \mathbf{t})} Y \sim N(\mu', \Sigma') = N(\mathbf{A}\mu + \mathbf{t}, \mathbf{A}\Sigma\mathbf{A}^T)$. A sűrűségfüggvények $N(\mu, \Sigma)$ és $N(\mu', \Sigma')$ paraméterei az input képekről egyszerűen számolhatók. A p és s között az alábbi

kapcsolat áll fent [5; 6; 7]

$$s(\mathbf{y}) = \frac{1}{2\pi\sqrt{|\Sigma'|}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mu')^T \Sigma'^{-1}(\mathbf{y} - \mu')\right) = \frac{1}{|\mathbf{A}|} p(\mathbf{x}) ,$$

ahol $|\mathbf{A}|$ könnyedén adódik az $|\mathbf{A}| = \sqrt{|\Sigma'|/|\Sigma|}$ összefüggésből, mivel $\mathbf{A}\Sigma\mathbf{A}^T = \Sigma'$. Néhány szükséges ekvivalens átalakítás után a *Mahalanobis távolságot* kapjuk meg [5; 6; 7]. Ezután a következőképpen definiálhatjuk a $P, S : \mathbb{R}^n \rightarrow \mathbb{R}$ kovariáns függvényeinket [5; 6; 7]

$$P(\mathbf{x}) = (\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu) \quad \text{és} \quad S(\mathbf{y}) = (\mathbf{y} - \mu')^T \Sigma'^{-1}(\mathbf{y} - \mu') .$$

Ekkor

$$P(\mathbf{x}) = S(\mathbf{A}\mathbf{x} + \mathbf{t}) = S(\mathbf{y}) . \quad (12)$$

Összeszorozva egymással a (11) és (12) egyenleteket, az összes pontmegfeleltetést összeintegrálni kapjuk, hogy [5; 6; 7]

$$\int_{\mathcal{F}_t} \mathbf{x} P(\mathbf{x}) \, d\mathbf{x} = |\mathbf{A}|^{-1} \int_{\mathcal{F}_o} \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t}) S(\mathbf{y}) \, d\mathbf{y} ,$$

ahol az $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t})$, $d\mathbf{x} = |\mathbf{A}|^{-1} d\mathbf{y}$ integrál transzformációt alkalmaztuk. További lineárisan független egyenletek generálásához nemlineáris $\omega : \mathbb{R} \rightarrow \mathbb{R}$ függvényeket használunk, melyek a (4) egyenlet alapján új egyenleteket generálnak [5; 6; 7]

$$\int_{\mathcal{F}_t} \mathbf{x} \omega(P(\mathbf{x})) \, d\mathbf{x} = |\mathbf{A}|^{-1} \int_{\mathcal{F}_o} \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t}) \omega(S(\mathbf{y})) \, d\mathbf{y} . \quad (13)$$

Egy lineárisan független $\{\omega_i\}_{i=1}^m$ halmazt felhasználva a fenti integrálokat az alábbi lineáris egyenletrendszerek segítségével is kifejezhetjük bármely $k = 1, \dots, n$ -re [5; 6; 7]

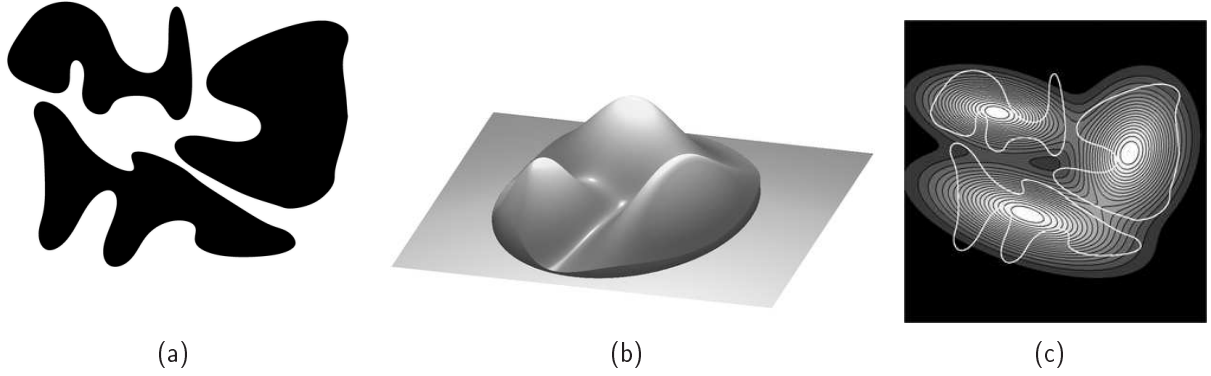
$$|\mathbf{A}| \int_{\mathcal{F}_t} x_k \omega_j(P(\mathbf{x})) \, d\mathbf{x} = \sum_{i=1}^n q_{ki} \int_{\mathcal{F}_o} y_i \omega_j(S(\mathbf{y})) \, d\mathbf{y} + q_{k(n+1)} \int_{\mathcal{F}_o} \omega_j(S(\mathbf{y})) \, d\mathbf{y} ,$$

ahol $j = 1, \dots, m$. Ezen lineáris egyenletrendszer megoldása határozza meg a transzformáció paramétereit. Ha $m > 3$, akkor az egyenletrendszer túlhatározottá válik, és az eredményt legkisebb négyzetes értelemben vett megoldással kapjuk.

Kovariáns függvénypárjainkat definiálhatjuk úgy, hogy $\mathcal{P}, \mathcal{S} : \mathbb{R}^n \rightarrow \mathbb{R}$ [6; 7]

$$\mathcal{P}(\mathbf{x}) = 2\pi\sqrt{|\Sigma|} p(\mathbf{x}) \quad \text{és} \quad \mathcal{S}(\mathbf{y}) = 2\pi\sqrt{|\Sigma'|} s(\mathbf{y}) . \quad (14)$$

Tegyük fel, hogy a *sablon* képen lévő objektum $\ell \geq 2$ diszjunkt részt tartalmaz. Ekkor mindegyik résznek pontosan egy alakzat felel meg a *megfigyelésen*. Következésképpen *kovariáns*



2. ábra. Gauss sűrűségfüggvények illesztése összetett alakzatok felett konzisztens kiszínezést ad. (a) Eredeti alakzat. (b) Gauss sűrűségfüggvény 3D-s ábrája az $r = 2$ által meghatározott ellipszis felett. (c) A Gauss sűrűségfüggvények szürkeárnyaltos képként reprezentálva. A fehér vonal az egyes komponensek határát jelöli.

függvénypárjainkat az összes alakzatra külön-külön is definiálhatjuk úgy, hogy $\mathcal{P}_i(\mathbf{x})$, $\mathcal{S}_i(\mathbf{y})$ kielégíti a (12) egyenletet. Továbbá az egész objektum saját maga is meghatároz egy *kovariáns függvénypárt*. Mivel az objektumok közötti megfeleltetés általában nem ismert, összegezzük ezeket az összefüggéseket, mely az alábbi *kovariáns függvényekhez* vezet [6; 7]:

$$P(\mathbf{x}) \equiv \sum_{i=0}^m \mathcal{P}_i(\mathbf{x}) = \sum_{i=0}^m \mathcal{S}_i(\mathbf{y}) \equiv S(\mathbf{y}). \quad (15)$$

Megjegyezzük, hogy ezen normalizálatlan Gauss sűrűségfüggvények keveréke tekinthető úgy, mint a *sablon* és a *megfigyelés* képek konzisztens kiszínezése (lásd 2. ábra), mely megőrzi a transzformáció hatását. A (13) egyenlethez hasonlóan az alábbi összefüggést is használhatjuk [6; 7]:

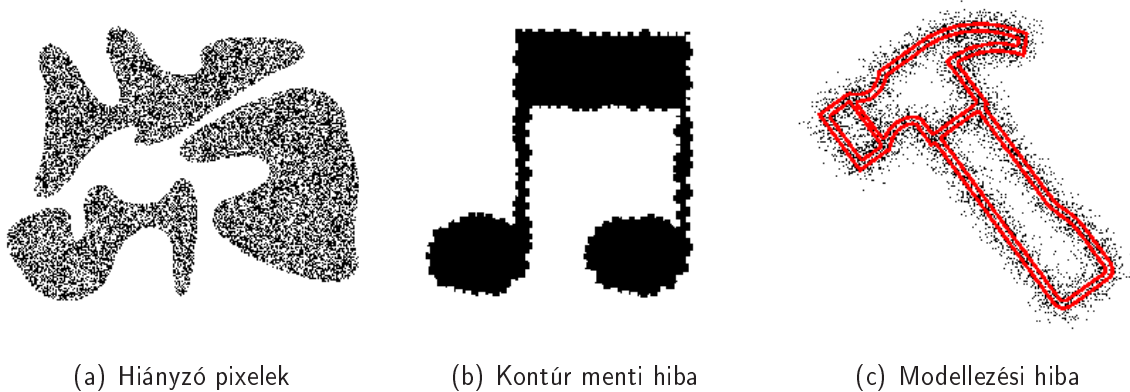
$$\sum_{i=1}^{\ell} \int_{\mathcal{F}_t} \mathbf{x} \omega(\mathcal{P}_i(\mathbf{x})) \, d\mathbf{x} = |\mathbf{A}|^{-1} \int_{\mathcal{F}_o} \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t}) \sum_{i=1}^{\ell} \omega(\mathcal{S}_i(\mathbf{y})) \, d\mathbf{y}.$$

A (13) egyenletében lévő integrálási tartományok megválasztására az \mathcal{F}_t és \mathcal{F}_o előtérpíxelek halmaza egy elég nyilvánvaló lehetőség [5]. Egyértelmű, hogy ennek az a hátránya, hogy bármilyen szegmentálási hiba rossz, hibás integrálokat okoz, mely rossz illesztéshez vezet. Az itt alkalmazott alapötlet, hogy a teljes *sablon* és *megfigyelés* kép statisztikáját használjuk fel. Mivel a megadott sűrűségfüggvények szintvonalai ellipszisek, természetes, hogy ezeket az ellipsziseket válasszuk az egymásnak megfelelő integrálási tartományoknak.

Két különböző megközelítést is javasolunk a lineáris egyenletrendszer felírásához. Először, mikor az objektum egy részből áll, csak az előtérpontok halmazát tudjuk felhasználni integrálási tartományként. Ebben az esetben a *kovariáns függvénypárunkat* a Mahalanobis távolság

segítségével adjuk meg, melyet az alakzatok definiálnak, és ekkor az alakzat felett kell vennünk az integrálokat (Egysűrűségfüggvényes módszer). A második esetben, mikor összetett alakzatunk van, integrálási tartománynak ellipsziseket is választhatunk, illetve a *kovariáns függvényeinket* Gauss sűrűségfüggvények keveréke adja. Az egyenleteinket folytonos esetre vezettük le, a gyakorlatban viszont csak véges pontosságú digitális képek állnak rendelkezésre. Emiatt az integrálok véges összegek segítségével közelíthetők, melyet tetszőleges finomságú beosztás esetén számolhatunk (Többsűrűségfüggvényes módszer véges összegekkel). Mindazonáltal, mikor több részből álló alakzatok állnak rendelkezésre, egy numerikusan hatékony módszert mutatunk az integrálok kiszámítására, mely zárt képlettel kiszámolható (Többsűrűségfüggvényes módszer numerikusan hatékony számítással).

Bár az ω_i megválasztására más lehetőség is adódik, *hatványfüggvények* segítségével zárt képletet adhatunk az integrálok kiszámítására az \mathcal{F}_t és \mathcal{F}_o alapján számolt ellipszis tartományok felett. Ezen numerikus számításnak legnagyobb előnye a közel valós idejű hatékonyság. Tapasztalati úton azt találtuk, hogy a páratlan kitevőjű hatvány és gyökfüggvények azaz a $\{x, x^3, x^5, \sqrt[3]{x}, \sqrt[5]{x}\}$ halmaz megfelelő eredményt ad az összes tesztünk során.



3. ábra. Példák a robusztusság tesztelése során használt *megfigyelésekre*. c)-ben az eredeti alakzat kontúrja látható piros színnel megjelenítve.

Az algoritmus hatékonyságának elemzésére egy olyan képadatbázist hoztunk létre, mely szintetikusán transzformált *megfigyeléseket* tartalmaz, továbbá az alkalmazott transzformációk véletlenszerűen lettek kiválasztva. A módszert mind szintetikus, mind valós képeken teszteltük. Majd a (10) egyenletben definiált hibamértékeket értékeltük ki. A javasolt algoritmus robusztusságát hiányzó pixelek, kontúr menti hiba és modellezési hiba (lásd 3. ábra) esetén is elemeztük. Valós képek mellett, melyek önmagukban tartalmaznak ilyen hibákat, mesterségesen generált adatokon is teszteltünk. A 2. táblázat mutatja teszteredményekhez tartozó hibamértékek mediánjait.

A javasolt módszer sem pontmegfeleltetéseket, sem pedig bonyolult optimalizálási lépé-

Módszer	ϵ (pixel)	δ %	CPU idő (mp.)
Egyszerűségfüggvényes módszer [5]	0.64	0.31	0.48
Többszerűségfüggvényes módszer véges összegekkel [6]	0.58	0.25	4.65
Többszerűségfüggvényes módszer numerikusan hatékony számítással [7]	0.54	0.19	0.33

2. táblázat. Az algoritmus futtatása során kapott hibamértékek és a futási idők mediánjai 1500 véletlenszerűen kiválasztott *megfigyelés* esetén.

sek megoldását nem igényli. Lineáris időbonyolultságú és a deformáció mértékétől függetlenül megadja a pontos transzformációt. Összetett alakzatok esetén egy robusztus és numerikusan hatékony megoldást adtunk, mely közel valós idejű megoldás ad. A javasolt algoritmus lineáris időbonyolultságából adódóan nagy képeken is elég gyorsan fut, így nem kell kompromisszumot kötni a minőség rovására, ha a futási idő kritikus tényező. A teszteredmények mutatják, hogy a módszer jó illesztést biztosít mind valós, mind pedig szintetikus képek esetén. Továbbá a módszer robusztusságát is szemlétettük. Általánosságban azt mondhatjuk, hogy a módszerünk jó eredményt ad, amíg az alakzat első és másodrendű statisztikája nem változik meg drasztikusan, így minden olyan alkalmazás során jól használható, ahol a takarás minimális.

Széttört objektumok deformációjának helyreállítása

Most azt a problémát tekintjük, ahol egyszerre több különböző lineáris transzformációt illesztünk pontmegfeleltetések felhasználása nélkül, melyek egy globális nemlineáris deformációt eredményeznek egy eredeti objektum és annak széttört változata között. Célunk a *megfigyelés* $2 \leq \ell \in \mathbb{N}$ alakzatának helyreállítása az eredeti, a *sablon* képen lévő pocizójukba. A transzformáció nemlineáris, és ℓ darab lineáris transzformáció kompozíciójaként áll elő, ahol az i . transzformációt \mathbf{A}_i -vel jelöljük. Ez a feladat *puzzle* néven is ismert, mely nemcsak elméleti szempontból fontos, hanem számos alkalmazási területe is van, mint például az archeológia és az orvosi képfeldolgozás (pl. törött csontok helyreállítása).

Feltételezzük, hogy a *sablon* teljes szegmentálása nem ismert, azaz particionálása rejtett. Az input képeken az alakzatok címkézését a $\mathbb{L}_t, \mathbb{L}_o : \mathbb{P}^n \rightarrow \{0, 1, \dots, \ell\}$ függvények adják, melyek a 0 értéket a háttérhez rendelik. Pontosabban, mivel \mathbb{L}_t rejtett, a *sablon* kép particionálása is ismeretlen. Jelölje $\mathcal{D}_i = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{L}_t(\mathbf{x}) = i\}$ az i . alakzatot a *sablon* képet és $\mathcal{D}'_i = \{\mathbf{y} \in \mathbb{P}^n | \mathbb{L}_o(\mathbf{y}) = i\}$ annak transzformált változatát a *megfigyelés* képen. Ha az alakzatok közötti megfeleltetés ismert lenne, akkor bármely standard bináris regisztrációs módszert felhasználva, páronkénti illesztéssel a probléma megoldható lenne. Sajnos ezen megfeleltetések megtalálása nem könnyű feladat, emiatt *direkt megoldás* iránt érdeklődünk az objektumok

közötti megfeleltetések meghatározása nélkül.

Tekintsük most csak az i . alakzatot, ahol az (1) egyenlet úgy írható, hogy [4; 8]:

$$\mathbf{x} = \mathbf{A}_i \mathbf{y} .$$

Megjegyezzük, hogy ez akkor is fennáll, ha mindkét oldalra egy nemlineáris $\omega : \mathbb{P}^n \rightarrow \mathbb{R}$ függvényt alkalmazunk [4; 8]:

$$\omega(\mathbf{x}) = \omega(\mathbf{A}_i \mathbf{y}) . \quad (16)$$

\mathcal{D}_i feletti integrálással kapjuk, hogy [4; 8]

$$\int_{\mathcal{D}_i} \omega(\mathbf{x}) \, d\mathbf{x} = |\mathbf{A}_i| \int_{\mathcal{D}'_i} \omega(\mathbf{A}_i \mathbf{y}) \, d\mathbf{y} , \quad (17)$$

ahol az $\mathbf{x} = \mathbf{A}_i \mathbf{y}$, $d\mathbf{x} = |\mathbf{A}_i| d\mathbf{y}$ integrál transzformációt alkalmaztuk, és $|\mathbf{A}_i|$ az i . transzformáció Jacobi-féle determinánsa. A nemlineáris ω függvény közvetlenül a pontkoordinátákon hat, ezáltal közvetlenül az ismeretlen \mathbf{A}_i transzformáció paraméterein, és nemlineáris egyenletrendszerhez vezet [4; 8]. A (17) egyenlet alapján lineárisan független függvények alkalmazásával annyi egyenletet generálhatunk, amennyire szükség van.

Így a (17) egyenlet segítségével úgy sikerült összefüggéseket megadni az i . alakzatpárok között, hogy sem a *sablon* particionálása (\mathbb{L}_t rejtett címkézés) sem pedig az alakzatok közötti megfeleltetés nem ismert. Szokásos módszer összeadni az egyenleteket az összes \mathcal{D}_i tartományra vonatkozóan, majd úgy megoldani a problémát, hogy egyszerre, az összes paramétert keressük egyetlen egyenletrendszerben felírva. $\{\omega_j\}_{j=1}^m$ függvények egy halmazának felhasználásával és a (17) egyenlet alapján kapjuk, hogy [8]:

$$\sum_{i=1}^{\ell} \int_{\mathcal{D}_i} \omega_j(\mathbf{x}) \, d\mathbf{x} = \sum_{i=1}^{\ell} |\mathbf{A}_i| \int_{\mathcal{D}'_i} \omega_j(\mathbf{A}_i \mathbf{y}) \, d\mathbf{y} .$$

Legyen $\mathcal{F}_t := \cup_{i=1}^{\ell} \mathcal{D}_i$ a teljes *sablonnak* megfelelő tartomány, ahol $\mathcal{F}_t = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{L}_t(\mathbf{x}) \neq 0\}$. Így a fenti egyenlet bal és jobb oldala, az alábbi formában írható [8]

$$\sum_{i=1}^{\ell} \int_{\mathcal{D}_i} \omega_j(\mathbf{x}) \, d\mathbf{x} = \int_{\cup_{i=1}^{\ell} \mathcal{D}_i} \omega_j(\mathbf{x}) \, d\mathbf{x} = \int_{\mathcal{F}_t} \omega_j(\mathbf{x}) \, d\mathbf{x} ,$$

amely közvetlenül számolható az input képek alapján a \mathcal{D}_i particionálás ismerete nélkül. Így olyan egyenletrendszert kapunk, mely $\ell n(n+1)$ ismeretlent tartalmaz [8]:

$$\int_{\mathcal{F}_t} \omega_j(\mathbf{x}) \, d\mathbf{x} = \sum_{i=1}^{\ell} |\mathbf{A}_i| \int_{\mathcal{D}'_i} \omega_j(\mathbf{A}_i \mathbf{y}) \, d\mathbf{y} \quad j = 1, \dots, m . \quad (18)$$

Az egyenletrendszer megoldása a deformáció összes paraméterét megadja. Mivel bármely ω_j egy-egy egyenletet ad, így legalább $m \geq \ell n(n+1)$ lineárisan független függvény kell ℓ lineáris transzformációhoz. A gyakorlatban $m > \ell n(n+1)$, ami túlhatározott egyenletrendszerhez vezet, ahol az egyenletrendszert legkisebb négyzetes értelemben oldjuk meg.

Elméletileg bármely olyan függvény felhasználható a (18) egyenletrendszer előállításához, mely kielégíti a (16) egyenletet. A megoldást egy iteratív, legkisebb négyzetes minimalizáláson alapuló algoritmussal kapjuk (ilyen például a *Levenberg-Marquardt algoritmus*), mely jól megválasztott numerikus szerkezetet igényel. A egyenlet megoldónak minden iterációs lépésben ki kell értékelni az egyenleteket, emiatt olyan ω függvények alkalmazása a cél, melyek tisztán nemlineáris egyenletrendszerhez vezetnek, ahelyett, hogy integrál egyenleteket kapnánk (azaz az ismeretlenek nem szerepelnek az integrálokban). Korábban megmutattuk [4], hogy hatványfüggvények alkalmazásával polinom egyenletrendszert kapunk, ahol az integrálok előre kiszámolható konstansokká válnak. Ezt figyelembe véve, polinomok következő halmazát használjuk [8]

$$\{\omega_j\}_{j=1}^m = \{\mathbf{x} \mapsto x_1^{u_1} \dots x_n^{u_n} | u_1, \dots, u_n \in \mathbb{N}_0, \sum_{i=1}^n u_i \leq d\}, \quad (19)$$

ahol $\omega_j : \mathbb{P}^n \rightarrow \mathbb{R}$, illetve d a maximális fokszámot jelöli, és $m = \frac{1}{n!} \prod_{i=1}^n (d+i)$ pedig a polinomok számát adja meg.

Mindkét kép koordinátáit a $[-0.5, 0.5]^n$ hiperkockába *normalizáljuk* ezzel elkerülve a lehetséges numerikus hibákat, melyeket a magas kitevők okoznak. Legyen \mathbf{N}_t és \mathbf{N}_o a *sablon* és a *megfigyelés* képekhez tartozó normalizáló transzformáció. Mivel legkisebb négyzetes értelemben oldjuk meg a feladatot, így a (18) egyenlet algebrai hibájára vonatkozóan azt várjuk el, hogy minden egyenlet azonos hozzájárulással bírjon, ezzel garantálva a hiba kiegyensúlyozottságát. A gyakorlatban csupán véges pontosságú digitális képek állnak rendelkezésre, így az integrálokat csak az előtérpíxelek halmazán vett véges összegek segítségével, *közelítőleg* tudjuk megadni. Ezek viszont elhanyagolható hibát jelentenek a számítások során [8]

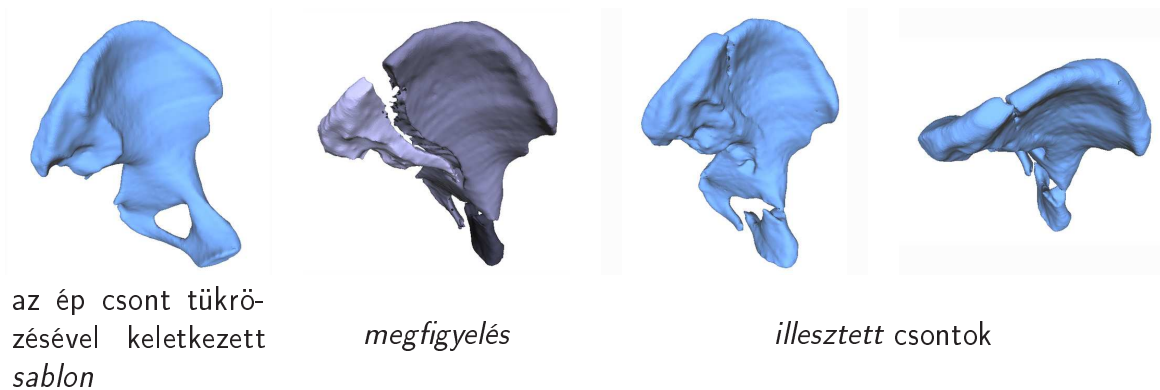
$$\frac{1}{c_j} \sum_{\mathbf{x} \in F_t} \omega_j(\mathbf{N}_t \mathbf{x}) = \frac{1}{c_j} \sum_{i=1}^{\ell} |\mathbf{A}_i| \sum_{\mathbf{y} \in D'_i} \omega_j(\mathbf{N}_o \mathbf{A}_i \mathbf{y}), \quad (20)$$

ahol c_j egy megfelelően választott konstans, melyet az origó középpontú és $\sqrt{n}/2$ sugarú hipergömbön vett $\int |\omega_j(\mathbf{x})| d\mathbf{x}$ integrál határoz meg.

A javasolt keretrendszert jól ismert lineáris transzformációkra, azaz 2D és 3D, affin és merev test transzformációkra alkalmaztuk. 2D affin transzformációkat gyakran használnak projektív torzulások lineáris közelítéseként. Ha egy objektum több részre törik, akkor az egyes

részek általában külön-külön, egy-egy merev test transzformációval torzulnak. A 3D merev test transzformáció pedig számos orvosi alkalmazás során játszik fontos szerepet.

A javasolt algoritmust egy nagyméretű szintetikus adatbázison teszteltük, mely 2D és 3D képeket tartalmaz, ahol lineáris (affin és merev test) transzformációkat alkalmaztunk. Ezután részletesen elemeztük a javasolt algoritmus numerikus stabilitását. Az eredmények mutatják, hogy a módszer szegmentálási hibákkal szemben robusztus. Az eredményeket 2D valós képeken és 3D-s műtéti tervezéssel kapcsolatos orvosi képeken mutattuk be. A gyakorlatban a szegmentálás nem eredményez tökéletes alakzatokat, ezért a javasolt megközelítés robusztuságát különböző típusú szegmentálási hibák esetén, illetve valós képeken (melyek önmagukban tartalmaznak ilyen hibákat) is bemutattuk.



4. ábra. Törött csontok helyreállítása.

Egy olyan újfajta keretrendszert mutattunk be, mely alakzatok összeillesztésére alkalmas, ahol a módszert 2D és 3D affin, illetve merev test transzformációkra alkalmaztuk (lásd 4. ábra). A klasszikus megoldásokkal szemben, melyek tulajdonságpontok kinyerésén és megfeleltetésén alapulnak, az itt javasolt megoldás minden további információ felhasználása nélkül keresi meg az illesztés paramétereit. Lényegében a módszer egy polinom egyenletrendszer felírása után, annak megoldásával közvetlenül megadja az ismeretlen transzformáció paramétereit. A 2D és 3D képeket tartalmazó adatbázisokon futtatott eredmények jól mutatják a módszer hatékonyságát és robusztusságát, továbbá a valós képeken kapott eredmények alapján azt mondhatjuk, hogy a módszernek több különböző alkalmazási területe is van. Legnagyobb előnye, hogy nincs szükség pontmegfeleltetésre, gyors és könnyen implementálható.

Az eredmények tézisszerű összefoglalása

A következőkben összefoglaljuk a disszertáció legfontosabb eredményeit. Az itt bemutatott eredmények több publikációban jelentek meg. A 3. táblázat mutatja, hogy az egyes tézispontokhoz, mely publikációk kapcsolódnak.

- I.) A Szerző egy olyan módszert mutat be, mely egy ismert alakzat és annak deformált változatai között lévő affin transzformációk paramétereinek becslésére alkalmas. Egy olyan újfajta megközelítést javasol, ahol a pontos transzformáció egy polinom egyenletrendszer megoldásával megkapható. A módszert szintetikus és valós képeken is tesztelte. A Szerző bemutatta a módszer robusztusságát szegmentálási hiba és additív geometriai zaj esetén, majd pedig azt, hogy a módszer sikeresen alkalmazható csípőprotézisről készült röntgen képek regisztrációjára.
- II.) A Szerző egy olyan újfajta megközelítést mutat be, mely egy ismert alakzat és annak deformált változatai között lévő affin transzformációk paramétereinek becslésére alkalmas. A Szerző megmutatja, hogyan kapható meg a pontos transzformáció az alakzatok felett megkonstruált Gauss sűrűségfüggvények (melyek megőrzik az ismeretlen transzformáció hatását) alapján felírt lineáris egyenletrendszer legkisebb négyzetes értelemben vett megoldásával. Több részből álló alakzatok esetén a Szerző egy robusztus és numerikusan hatékony megoldást mutat be, mellyel közel valós idejű hatékonyság érhető el. A módszert mind szintetikus, mind valós képeken tesztelte. A robusztusság mellett hiányzó pixelek, kontúr menti hiba és modellezési hiba esetén is szemléltette a módszer működését.
- III.) A Szerző széttört objektumok összeillesztésének problémáját vizsgálja pontmegfeleltetések felhasználása nélkül, ahol csak a teljes *sablon* kép szegmentálása ismert, az egyes alakzatoké nem. A Szerző a széttört objektumok között lineáris transzformációkat feltételez, és a módszert 2D és 3D affin transzformációk esetén mutatja be. Megmutatja, hogyan lehet megkonstruálni egy olyan polinom egyenletrendszert, melynek megoldása megadja az ismeretlen illesztési transzformáció paramétereit. A javasolt algoritmust nagyméretű szintetikus adatbázison tesztelte, mely 2D és 3D képeket is tartalmaz. A robusztusság és a módszer szegmentálási hibákra való érzékenysége mellett a módszer numerikus stabilitását is elemzi. Az eredményeket 2D valós képeken és egy 3D orvosi alkalmazás esetében (műtéti tervezés) is bemutatta.

	[3]	[4]	[5]	[6]	[7]	[8]
I.	•	•				
II.			•	•	•	
III.						•

3. táblázat. A tézispontokhoz kapcsolódó publikációk.

Hivatkozások

- [1] Rami Hagege and Joseph M. Francos. Parametric estimation of two dimensional affine transformations. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 305–308, Montreal, May 2004. IEEE.
- [2] Rami Hagege and Joseph M. Francos. Linear estimation of sequences of multi-dimensional affine transformations. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 785–788, Toulouse, France, May 2006. IEEE.
- [3] Csaba Domokos, Zoltan Kato, and Joseph M. Francos. Parametric estimation of affine deformations of binary images. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pages 889–892, Las Vegas, NV, USA, April 2008. IEEE.
- [4] Csaba Domokos and Zoltan Kato. Parametric estimation of affine deformations of planar shapes. *Pattern Recognition*, 43(3):569–578, March 2010.
- [5] Csaba Domokos and Zoltan Kato. Binary image registration using covariant Gaussian densities. In Aurélio Campilho and Mohamed Kamel, editors, *Proceedings of International Conference on Image Analysis and Recognition*, volume 5112 of *Lecture Notes in Computer Science*, pages 455–464, Póvoa de Varzim, Portugal, June 2008. Springer.
- [6] Csaba Domokos and Zoltan Kato. Affine alignment of compound objects: A direct approach. In *Proceedings of International Conference on Image Processing*, pages 169–172, Cairo, Egypt, November 2009. IEEE.
- [7] Csaba Domokos and Zoltan Kato. Affine shape alignment using covariant Gaussian densities: A direct solution. *IEEE Transactions on Image Processing*, 2010. Submitted.
- [8] Csaba Domokos and Zoltan Kato. Affine puzzle: Realigning deformed object fragments without correspondences. In Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, *Proceedings of European Conference on Computer Vision*, volume 6312 of *Lecture Notes in Computer Science*, pages 777–790, Heraklion, Crete, Greece, September 2010. Springer.